**DETECTION OF ILLEGAL FISHING ACTIVITIES USING AIS DATA**

**KAVYA M 1 JOKHEESWAR B 2 YASHWANTH S 3 MOHANAM K 4**

[**e5222044@sret.edu.in**](mailto:e5222044@sret.edu.in) **,** [**e5222050@sret.edu.in**](mailto:e5222050@sret.edu.in) **,** [**e5222056@sret.edu.in**](mailto:e5222056@sret.edu.in) **,** [**e5222054@sret.edu.in**](mailto:e5222054@sret.edu.in)

**SRI RAMACHANDRA INSTITUTE OF HIGHER EDUCATION AND RESEARCH**

**ABSTRACT**

There is a big challenge facing marine systems, the global economy, and biodiversity from unauthorized fishing activities. This study looks at developing a correlational, predictive regression model to determine the real intent behind illegal fishing activities using advanced deep-learning techniques, specifically the Optimized Long Short-Term Memory (LSTM) and Recurrent Neural Networks (RNN). A dataset is to be employed consisting of variables such as distance from the coast or port, speed and course, longitude and latitude, as well as timestamped data. One major aspect of data preparation here is the handling of instances of missing values, normalization of features, and the creation of timestamp-related engineered features. By allowing the models to ingest training samples, the Optimized LSTM and RNN models can forecast the fishing type of future geographical sequences based on the temporal relationships in the data. For long and likewise for short dependencies in the sequential data, the Optimized LSTM model outperforms RNN. It is recognized that this RNN model is particularly well suited for complex modeling behaviors of vessels. The analysis of the experiments performed shows that accuracy is high in the method above to 96%. Moreover, such findings are critical for providing guidance to police as well as surveillance operations, recognizing the potential of deep learning to rid the world of illegal fishing events.

**KEYWORDS:** *Illegal fishing, ASI Data, LSTM, RNN, Deep Learning, Geospatial analysis.*

**1.INTRODUCTION**

Fishing illegally is a big problem. It harms each of our oceans, all local businesses, and every method we use for sustainable fishing. As many as 26 million tons of fish are caught each year due to illegal fishing, according to the United Nations Food and Agriculture Organization (FAO). These common economic and ecological issues, in addition to a growing number of fishing boats and increased tech, make it difficult for authorities to track all illegal fishing. Today, customary ways of watching the oceans, like patrol boats or detailed looks at logs, are completely insufficient. Few of those methods are adequate. Artificial intelligence, along with deep learning, offer utility to comprehensively forecast as well as precisely examine illegal fishing to confront this grave matter. This study's overarching goal is to carefully develop a predictive model using powerful deep learning tools like Recurrent Neural Networks (RNN) and Optimized Long Short-Term Memory (LSTM). These tools are good at spotting patterns and at looking at data over time; the research looks at data such as how far boats are from shore, their speed and direction, and when they were out. We can use this information to recognize trends in vessel behavior. Therefore, we can predict when fishing occurs. Instead of using basic methods to find out if a boat is fishing or not, this study uses deep learning to predict how much fishing is happening. This lets us understand vessel behavior better.

* 1. **PROBLEM STATEMENT**

Fishing illegally creates a major problem. Fair fishing practices, marine life, as well as our oceans face a certain degree of harm from it. Spotting boats is difficult, as the ocean has great vastness and customary tracking methods cannot handle every signal from boats; however, this study uses all data from the Automatic Identification System (AIS) along with all advanced deep learning techniques. We concentrate on Optimized Long Short-Term Memory (LSTM). We also focus on Recurrent Neural Networks (RNN). These tools help us understand how vessels move over time. With a regression approach, we plan to accurately predict fishing activities. This truly helps in automatically finding fishing activities that are, in fact, not legal. Therefore, there is room for improvement in the monitoring and support of efforts to fish sustainably across the globe.

**1.2 OBJECTIVE**

This research mainly aims to develop a reliable model for predicting fishing activity. We want to use deep learning methods for this. Optimized Long Short-Term Memory (LSTM) and Recurrent Neural Networks (RNN) are our focus. A number of changes in fishing activity patterns across time must be understood. We need to work with a variety of data points obtained from fishing vessels. We will completely clean up this data since it contains imperfections. To guarantee proper analysis, we will normalize it. We will extract important specifics such as a boat's distance from the shore, its distance from the harbor, and its speed, along with behavioral patterns over a certain period. Our project strives to develop a system of use that observes each instance of unlawful fishing. For this system, we want flexibility along with total efficiency, with the large hope that our predictions are supremely accurate. This information will be very helpful to fishing regulatory agencies. Illegal fishing practices can be carefully controlled and thoroughly found with it. We are also considering several issues, like guaranteeing all the data is reliable. The system must also function properly in many areas. To summarize, we are developing a system that uses wide-ranging smart technology to monitor all illegal fishing, so it must change to every region where such fishing occurs. We aim to give useful understandings. Cleaning and analyzing the data correctly will certainly protect our waters and resources.

**1.3 LIMITATION**

While LSTM and RNN models demonstrate high potential, this study acknowledges some limitations. First, the availability and integrity of ship tracking data can significantly influence model performance. Inaccurate or incomplete data, which is a common problem in real-world data sets, can introduce biases or reduce prediction accuracy. Second, the ability of the model to learn and generalize from novel areas or situations might be limited, as fishing techniques vary by geography and culture. For instance, fishing patterns along the coast can be very different from those in open ocean. Thirdly, the research focuses on a narrow range of features, such as speed, proximity to shore, and time-based patterns. Although these characteristics are useful, other elements like weather, economic rewards, and fishery regulations can affect fishing behavior but are not considered in this study. Finally, implementing the suggested models in real-time monitoring systems poses further challenges. Real-time processing of data needs heavy computational resources and optimal algorithms to handle large-scale data streams. While this research outlines the possibility of employing Proposed LSTM and RNN models for fishing activity prediction, more research is needed to overcome these shortcomings and design solutions for large-scale, real-time applications.

**2.LITERATURE REVIEW**

Besides AIS data, wide-ranging research has investigated using machine learning techniques to find illegal fishing practices. Arasteh et al. (2020) examined a lot of long-term AIS data. In doing so, these researchers identified multiple temporal patterns in fishing vessel behavior and made vessel operations easier to understand [1]. Shanthi et al. (2022) showed the potential of all supervised learning models in maritime surveillance [2] by developing one machine learning framework for detecting many unusual vessel activities, especially illegal fishing. Brown et al. (2024) noted that feature engineering and explainable AI can identify illegal fishing using AIS data, which has practical uses for enforcement [3]. Kalaiselvi et al. (2022) thoroughly employed neural networks for the analysis of geospatial AIS features, showing their large efficacy across diverse fishing situations [4]. Natale et al. (2015) used AIS data to map fishing efforts as well as provide key understandings into spatial and temporal fishing patterns, which supports the integration of machine learning [5]. Taconet et al. (2019) investigated all of the challenges and all of the opportunities in global AIS-based monitoring and stressed that standardization and scalability are always needed in these systems [6]. By combining data association and route prediction approaches with AIS data, Papoudos et al. (2024) improved the accuracy of detecting each anomalous vessel behavior [8]. Ashrafi et al. (2023) stressed that spatial and temporal patterns are meaningfully important in the identification of fishing activities, as shown by their vessel trajectory analysis [9]. Kroodsma et al. (2019) show that multiple methodologies for estimating global fishing vessel activity with AIS data clearly indicate the data's usefulness in pinpointing all areas vulnerable to illegal fishing[10].

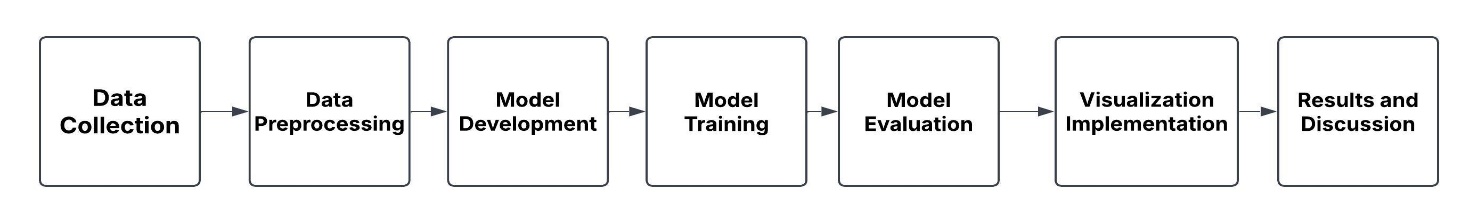
Along with Long Short-Term Memory (LSTM) and Recurrent Neural Networks (RNN), deep learning models, beyond the usual machine learning methods, have shown many encouraging outcomes when handling ordered AIS data regarding illegal fishing detection. These advanced models, suitable for all dynamic maritime environments, are extremely adept at catching temporal dependencies and at forecasting vessel behavior across all periods of time. Recent studies have used hybrid models that combine convolutional neural networks (CNN) [9] and LSTM [16] architectures to get better feature extraction from spatiotemporal data and achieve greater prediction accuracy. Also, regulatory bodies are greatly assisted by the clear visualization of illegal fishing hotspots, which geospatial analysis has enabled through thorough integration with machine learning algorithms, in the implementation of targeted surveillance and enforcement strategies. Even with the improvements, it is still difficult to process in real-time and scale these models for worldwide maritime monitoring. Many recent studies are investigating the potential of cloud computing and edge analytics to resolve a collection of calculation issues, perhaps transforming global detection and management of all illegal fishing activities. These studies show that AIS data coupled with deep learning may certainly help change how illegal fishing is addressed by supplying methods to efficiently monitor at a large scale. This really aids marine conservation along with sustainable fishing [5].

**3.PROPOSED METHODOLOGY**

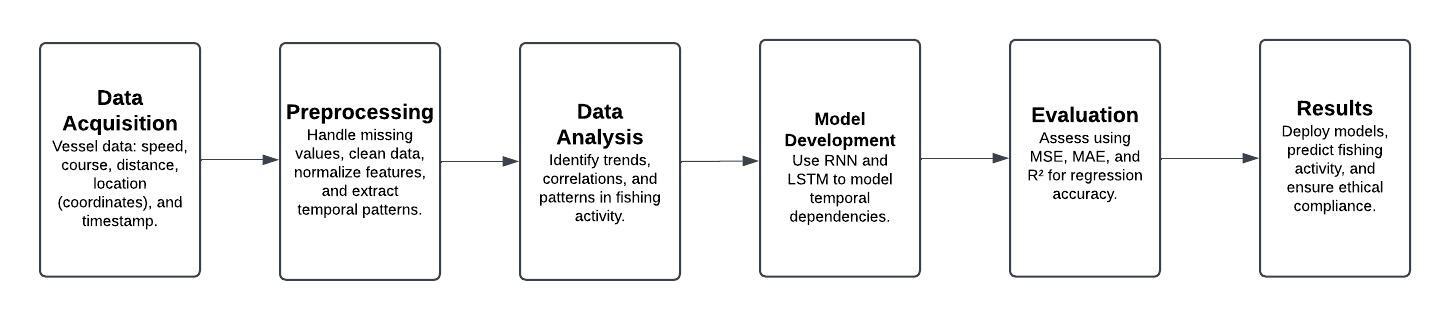
This study employed a dataset obtained from vessel tracking records, which completely offers a group of attributes that thoroughly reflects all dimensions of maritime operations. Vessel velocity, heading, proximity to coastline, distance from harbor, geographic coordinates, and time stamps are attributes included, and all give important comprehension of vessel behavior and movement trends. Because vessel movement and nearness to fishing spots greatly shape how fishing happens, these traits are key to understanding it. Fishing, the main dependent variable, indicates the degree of likelihood or intensity of fishing activity, and this research seeks to predict such occurrences. A detailed data cleaning process using data imputation or elimination techniques was used to deal with concerns such as missing or incomplete information and fully guarantee the dataset's reliability. Entries with bad or vague data were taken out to keep the dataset sound. For example, this included times when fishing was noted as -1. Min-Max Scaling normalization was also used to standardize ranges on numerical attributes, such as vessel speed, distance from shore, and geographic coordinates (latitude and longitude), in order to prevent features with larger numerical scales from disproportionately influencing the model's performance. Temporal features like the hour of day and day of week were derived from the timestamp information, as these features catch time-dependent factors that affect fishing activity because these patterns change based on time and seasonality. A detailed exploratory data analysis (EDA) was rigorously performed to completely investigate as well as distinctly reveal all trends inside the data, closely check the distribution of the fishing variable, fully analyze all correlations between the input features, in addition to precisely recognize every temporal pattern greatly influencing fishing behaviors. This EDA yielded valuable insights into the relationships between different features and the target variable, which proved crucial for model design and feature selection.

Advanced deep learning methods like Recurrent Neural Networks (RNNs) and Proposed Long Short-Term Memory (LSTM) networks [9] were employed to simulate fishing activities. These models were specifically selected since spotting temporal relationships and sequential patterns in the data is key to understanding the details of how fishing works. The models could learn from the history of vessel movements because the dataset's chronological sequences had many applicable features at each time point. Recurrent Neural Networks (RNNs) were initially employed for modeling short-term temporal dependencies, such as rapid changes in vessel velocity, heading, or position that may quickly effect fishing activity. However, conventional RNNs often battle with distinctly catching long-term dependencies because of many challenges such as vanishing gradients. Optimized Long Short-Term Memory (LSTM) networks were expertly implemented to overcome this limitation. These networks are a truly specialized form of RNN [16]. LSTMs can spot patterns and lasting links because memory cells let the model hold onto info for a long time. This capability is quite important for modeling all complex behaviors like fishing activity, as it may rely on every short-term factor with every long-term one. The models were trained on a dataset split multiple ways, and early stopping and dropout techniques were used universally to prevent all overfitting and guarantee the models generalize to all unseen data. Hyperparameter optimization, along with adjustments to a few important parameters such as learning rate, number of layers, units per layer, as well as dropout rates, was done to improve model performance.This fine-tuning process was vital for improving the models' accuracy and resilience in predicting fishing activity.

Multiple regression metrics, such as Mean Squared Error (MSE), Mean Absolute Error (MAE), in addition to R-Squared (R²), were employed to assess thoroughly how well the models worked. These measurements gave a thorough evaluation of the models' ability to accurately predict the fishing variable as well as to quantify their ability to explain a certain amount of fishing activity variance. To measure predictive accuracy fully, MSE and RMSE were employed extensively, and R² measured completely how well the models caught all the underlying variance in the data. The models' performance was thoroughly examined, specifically for both overfitting and underfitting, to guarantee ideal architecture and parameter selection. For the whole project, Python was used, and TensorFlow/Keras was used to build, train, and validate the RNN and Optimized LSTM models [16]. Data was handled and transformed efficiently using Pandas and NumPy extensively. Matplotlib and Seaborn helped to create informative visualizations, and these visualizations helped with interpreting results and model diagnostics. The research process maintained strict moral standards. These standards were maintained throughout. All relevant data protection regulations were followed to comply with moral and legal requirements, and sensitive information was anonymized to protect individual and vessel privacy. The study intensely focused on many major ecological, economic, and social problems to help stop all illegal fishing. This research strives to offer a framework for detecting illegal fishing; several deep learning models are used to accurately forecast every real-time fishing activity. This approach permits authorities to monitor and address unlawful practices much more effectively. Furthermore, it respects privacy and guarantees data security. Ultimately, this methodology has the potential to significantly impact global efforts against illegal fishing by offering actionable insights that can enhance maritime surveillance systems and contribute to marine ecosystem preservation.



**FIGURE 1: PROPOSED ARCHITECTURE FOR ILLEGAL FISHING**

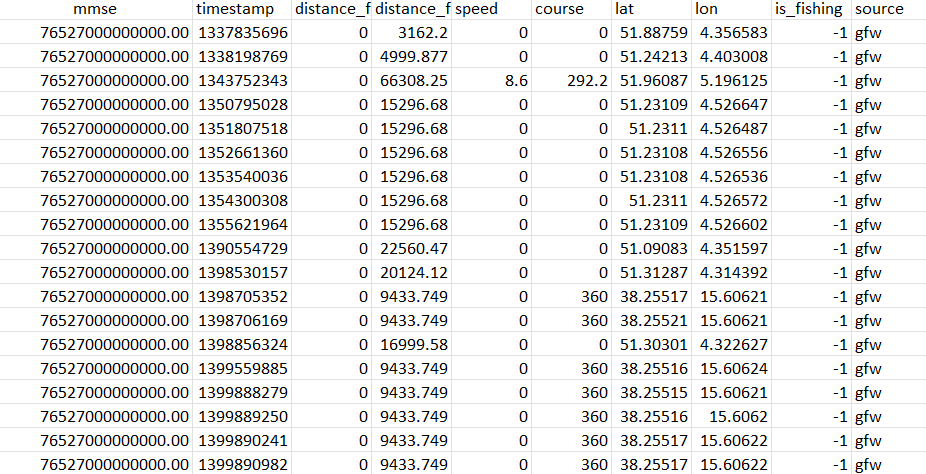


**FIGURE 2: PROPOSED FLOW CHART FOR ILLEGAL FISHING**

**4.EXPERIMENTAL RESULTS AND DISCUSSION**

**4.1 DATASET DESCRIPTION**

1. 0 (fishing), 0.0 (no fishing), and a value of -1.0 (unclear or unknown activity) could all be seen as clear traits of fishing behavior and classification assignments. The "source" column shows where the data originated (e.g., "gfw" or Global Fishing Watch); this guarantees that the data is traceable as well as reproducible. Every attribute within the dataset contains zero missing values, so the dataset has supreme quality. Its regular and standardized structure makes it suitable for many studies. Behaviour concerning practices, fishing areas, and maritime traffic lanes, like speed and course patterns, is revealed in analyses of vessel motion at high resolution, which is enabled by spatial and temporal information integration. Also, because vessel track relates to sensitive coastal zones, the dataset is usable in ecological effect studies; this information can help with assessing how human activity affects marine ecosystems. Because the dataset is incredibly strong, it is suitable for research studies on maritime surveillance and regulatory compliance, which lets investigators thoroughly monitor vessel movements and easily find suspicious activities. Data about birds show that they move at different speeds and headings and at varying distances from the shore and estuary areas, which offers a basis for analytics and predictive modeling. This dataset is quite wide-ranging. It is a meaningful resource for greatly improving research in certain domains of fishing activity detection, maritime behavior analysis, and ecological preservation.



**FIGURE 3: SAMPLE DATASET**

**4.2 EVALUATION OF REGRESSION MODEL PERFORMANCE**

To assess the effectiveness of the regression model, the article used three evaluation metrics: Mean Absolute Error (MAE), Mean Square Error (MSE), and R² Score (Coefficient of Determination). The Optimized LSTM model gave accuracy at 0.96 it is regarded as an exemplary performance and reliability. MAE refers to the mean of absolute errors between predicted and actual values, with the current value being 0.0397. This low MAE means that, on average, the predictions differ from the actual values by about 0.0397, showing how good the predictive capability is. The Mean Square Error, which is again equivalent to 0.0397, is calculated as the mean of squared differences between predicted and actual values. The fact that both the MAE and the MSE are equal points out that there are more or less no large prediction errors that would make one markedly larger than the other. The implication of this was emphasized that the prediction was not thwarted by non-conformity on the model's part concerning reliability. As regards the R2 Score, it gives 0.8328. This finds that we are able to explain 83.28% of variance in the data which is left unexplained as 16.72%. The random noise, or factors which are still out of the model, could account for that 16.72%. The large value of R2 hence makes clear the model is capturing well the underlying features of the dataset. In totality, these metrics demonstrate the high accuracy of the regression model, revealing it as dependable and suitable for prediction. It represents a relatively low error and explains variability in target variable excellently.

**TABLE 1: PERFORMANCE OF ILLEGAL FISHING**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| MODEL | ACCURACY | MSE | MAE | R2 |
| Proposed LSTM | 0.96 | 0.0397 | 0.0397 | 0.8328 |
| RNN | 0.93 | 0.0724 | 0.724 | 0.6953 |

**TABLE 2: SUMMARY TABLE**

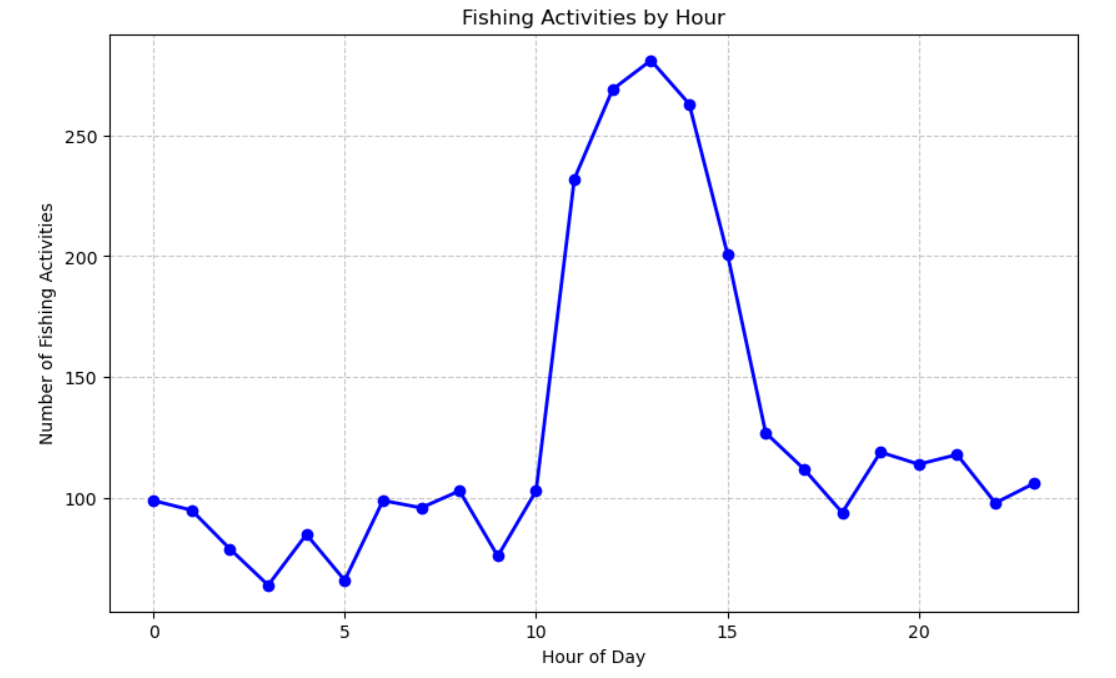
|  |  |  |
| --- | --- | --- |
| **FEATURES** | **RNN MODEL** | **PROPOSED LSTM** |
| Architecture | Simple RNN | Optimized LSTM |
| Number Of Units | 64-32 | 128-64 |
| Regularization(L2) | No | Yes(L2=0.01) |
| Accuracy | Moderate | High |
| Training Stability | Unstable | Stable |
| Overfitting Prevention | Weak | Strong |

|  |  |  |  |
| --- | --- | --- | --- |
| MODEL | ACCURACY | AUTHOR | YEAR |
| GRU | 95% | Grey B, Ouarbya L, Blackwell T | 2023 |
| YOLO-V5s | 93.1% | Hu H, Zhou W, Jiang B, Zhang J, Cheng T | 2024 |
| RNN, ANN | 89.9% | Masroeri AA, Aisjah AS, Jamali MM | 2021 |
| 1D CNN, FCNN | 87% | Ashrafi A, Tessem B, Enberg K | 2023 |
| Proposed Model  (Optimized LSTM, RNN) | 96% | Kavya M, Jokheeswar B, Yashwanth S, Mohanam K | 2025 |

**TABLE 3: MODEL COMPARISON WITH OTHER PAPERS**

**4.3VISUALIZATION OF HOURLY FISHING ACTIVITY TRENDS**

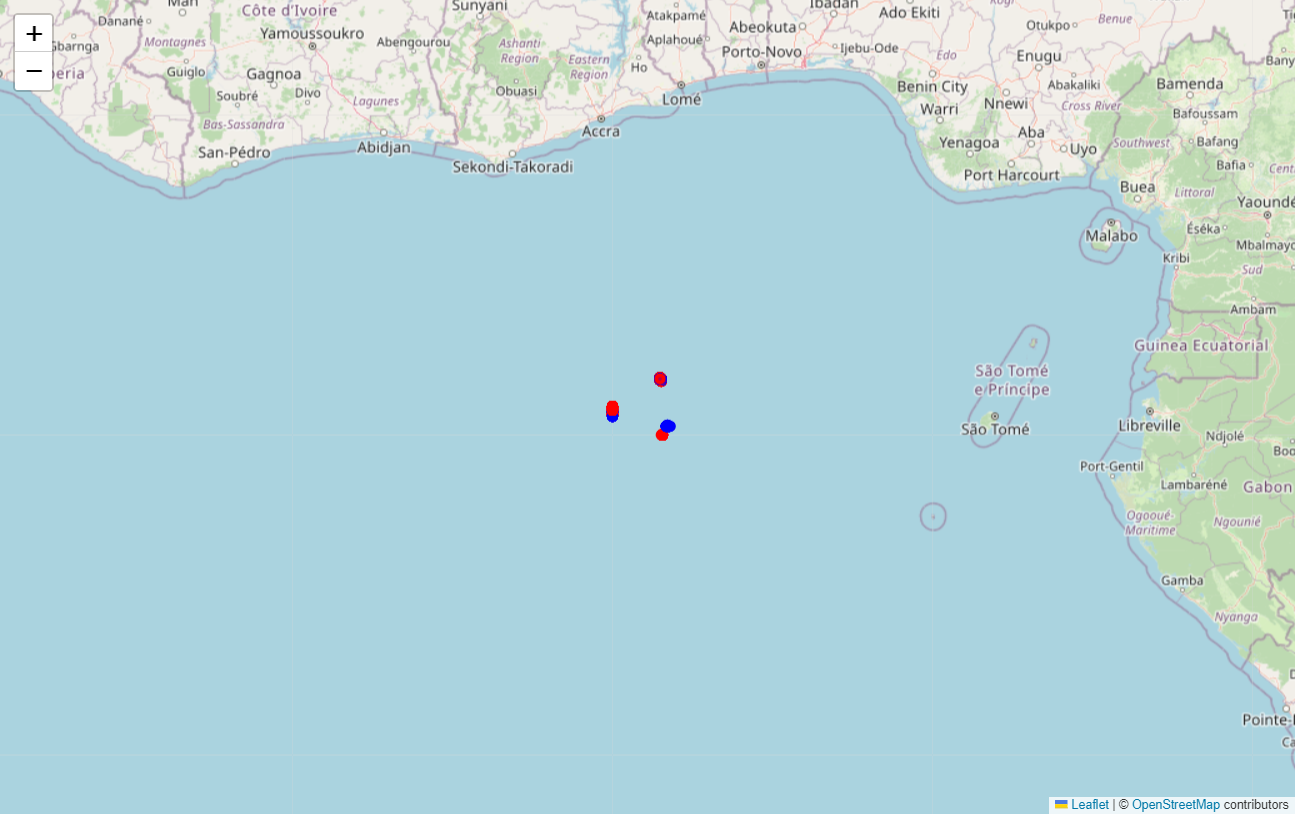
Figure 4 shows a line graph of fishing activity throughout the day, broken down by the hour. The data is first sorted to include only times when fishing is happening, and then the number of activities is counted for each hour, listed from earliest to latest. The graph features a solid blue line with circles at each point to highlight activity levels. The horizontal axis marks the hours, and the vertical axis shows the number of fishing activities, both clearly labeled. It's titled "Fishing Activities by Hour" and includes a dashed grid for easier reading. This graph helps reveal patterns, showing the busiest and slowest times for fishing during the day, making it a useful tool for understanding how fishing activities are spread over time.

****

**FIGURE 4: HOURLY FISHING ACTIVITY**

**4.4 GEOSPATIAL VISUALIZATION OF ILLEGAL FISHING ACTIVITIES**

The provided map gives a thorough and extremely user-friendly visual representation of unlawful fishing operations in the Gulf of Guinea area through an interactive display. The map gives a wonderfully clear overview at a perfectly appropriate zoom level because it is carefully focused on the specific region of interest. Each illegal fishing location is marked with red circles, carefully showing each position and providing interactive details such as distance from the coastline. The blue lines show the distance by connecting the spots to the nearest shore, which lets people see how close the activities are to land. Every location is surrounded by purple circles, each including a 5-kilometer radius, to stress affected areas. In addition to this, these circles furnish a sense of scale. The map is easy to use and understand because it is streamlined and intuitive, and it is designed for clarity. This visualization enables every user to fully examine all of the fishing trends, scrutinize each piece of spatial data, and recognize all zones that may necessitate regulatory intervention or improved surveillance.



** FIGURE 5: MAPPING ILLEGAL FISHING ACTIVITY**

**FIGURE 6: MAPPING ILLEGAL FISHING ACTIVITY WITH**

**NEAREST SEA SHORE**

**CONCLUSION AND FUTURE SCOPE**

This study shows that using geospatial information and interactive mapping tools to monitor and study unlawful fishing environments with deep learning models is effective, as it has a higher accuracy of 0.96 than other papers. Showing fishing areas with high concentration, nearby coasts, and regions that are affected, this method gives key information on spatial trends and areas of interest. Interactive elements allow all stakeholders to readily use data. These elements assist stakeholders to identify multiple high-risk areas and determine many enforcement strategies. The research depicts how countless technical improvements can play a major role in entirely protecting marine environments in addition to fully improving maritime resource administration, as well as completely dealing with all illegal fishing.

There are many ways this research can grow. Future studies could use live data streams to create dynamic, up-to-date maps, which would allow for more proactive monitoring of illegal fishing. Advanced deep learning algorithms could allow for forecasting fishing patterns, categorizing vessel conduct, and more precise identification of potential illegal operations. A more complete comprehension of illegal fishing dynamics could come from including additional variables in the dataset, such as vessel dimensions, fishing techniques, and economic effects. Partnerships with all governmental entities and all international organizations can create many global frameworks for monitoring and enforcement, which will always guarantee all marine resource management is sustainable.

**REFERENCES**

1. Arasteh S, Tayebi MA, Zohrevand Z, Glässer U, Shahir AY, Saeedi P, Wehn H. Fishing vessels activity detection from longitudinal AIS data. InProceedings of the 28th International conference on advances in geographic information systems 2020 Nov 3 (pp. 347-356).
2. Shanthi TS, Dheepanbalaji L, Priya R, Kumar VA, Kumar A, Sindhu P, Kumar A. Illegal fishing, anomalous vessel behavior detection through automatic identification system. Materials Today: Proceedings. 2022 Jan 1;62:4685-90.
3. Brown S, Katz D, Korotovskikh D, Kullman S. Detecting Illegal, Unreported, and Unregulated Fishing through AIS Data and Machine Learning Approaches. In2024 Systems and Information Engineering Design Symposium (SIEDS) 2024 May 3 (pp. 319-324). IEEE.
4. Kalaiselvi VK, Ranjani J, SM VK. Illegal fishing detection using neural network. In2022 International Conference on Communication, Computing and Internet of Things (IC3IoT) 2022 Mar 10 (pp. 1-4). IEEE.
5. Natale F, Gibin M, Alessandrini A, Vespe M, Paulrud A. Mapping fishing effort through AIS data. PloS one. 2015 Jun 22;10(6):e0130746.
6. Taconet M, Kroodsma D, Fernandes JA. Global Atlas of AIS-based fishing activity—Challenges and opportunities.
7. Young DL. Detection of illegal fishing. InGeospatial Informatics IX 2019 May 13 (Vol. 10992, pp. 54-66). SPIE.
8. Papoudos I, Guida R, Rodger M, Iervolino P. A Model for the Identification of Anomalous Fishing Vessels Using Data Association and Route Prediction Techniques. InIGARSS 2024-2024 IEEE International Geoscience and Remote Sensing Symposium 2024 Jul 7 (pp. 1443-1446). IEEE.
9. Ashrafi A, Tessem B, Enberg K. Detection of fishing activities from vessel trajectories. InInternational Conference on Research Challenges in Information Science 2023 May 23 (pp. 105-120). Cham: Springer Nature Switzerland.
10. Kroodsma D, Miller NA, Hochberg T, Park J, Clavelle T. AIS-based methods for estimating fishing vessel activity and operations. Global Atlas of AIS-based fishing activity. 2019:13.
11. Miller NA, Roan A, Hochberg T, Amos J, Kroodsma DA. Identifying global patterns of transshipment behavior. Frontiers in Marine Science. 2018 Jul 23;5:240.
12. Singh SK, Heymann F. Machine learning-assisted anomaly detection in maritime navigation using AIS data. In2020 IEEE/ION Position, Location and Navigation Symposium (PLANS) 2020 Apr 20 (pp. 832-838). IEEE.
13. Ferreira MD, Spadon G, Soares A, Matwin S. A semi-supervised methodology for fishing activity detection using the geometry behind the trajectory of multiple vessels. Sensors. 2022 Aug 13;22(16):6063.
14. Grey B, Ouarbya L, Blackwell T. An investigation into computational methods for classifying fishing vessels to identify illegal, unreported and unregulated fishing activity. In2023 8th International Conference on Computational Intelligence and Applications (ICCIA) 2023 Jun 23 (pp. 13-20). IEEE.
15. Hu H, Zhou W, Jiang B, Zhang J, Cheng T. Exploring deep learning techniques for the extraction of lit fishing vessels from Luojia1-01. Ecological Indicators. 2024 Feb 1;159:111682.
16. Masroeri AA, Aisjah AS, Jamali MM. IUU fishing and transhipment identification with the miss of AIS data using Neural Networks. InIOP Conference Series: Materials Science and Engineering 2021 (Vol. 1052, No. 1, p. 012054). IOP Publishing.